

Phase transition in random distance graphs on the torus

Fioralba Ajazi^{1, 2}, George M. Napolitano¹, and Tatyana Turova^{1, 3}

¹*Department of Mathematical Statistics, Faculty of Science, Lund University, Sölvegatan 18, 22100, Lund, Sweden.*

²*Department of Marketing and Operations Management, Faculty of Business and Economics, University of Lausanne, CH-1015, Switzerland.*

³*Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Institutskaja str., 142290, Pushchino, Moscow Region, Russia.*

Abstract

We apply here methods of inhomogeneous random graphs to a class of random distance graphs. This provides an example outside of the rank 1 models which is still solvable as long as the largest connected component is concerned. In particular, we show that some random distance graphs behave exactly as the classical Erdős-Rényi model not only in the supercritical phase (as was already known) but in the subcritical case as well.

1 Introduction

Random distance graphs, that is random graphs where the vertices are in a metric space, and the connection probabilities depend on the distance between the vertices, form a particular subclass of a general inhomogeneous random graph models [6]. The theory of inhomogeneous random graphs was founded in the seminal paper [6], where the authors gave a common ground for numerous models introduced and studied previously almost independently.

The graphs treated in [6] have the following common feature: the edges are independent and the probability of edges, is, roughly speaking, of order $1/n$, where n is the size of the graph. Shortly, each out of n vertices is assigned a type, i.e., a value in some separable metric space \mathcal{S} . Given a set

of such values $\{x_1, \dots, x_n\}$ any two vertices i and j are connected with a probability

$$p_n(i, j) = \min \left\{ \frac{\kappa(x_i, x_j)}{n}, 1 \right\}, \quad (1)$$

where κ is a symmetric nonnegative measurable function.

Random distance graphs assume usually that the vertices are in \mathbb{R}^d or \mathbb{Z}^d . Hence, their coordinates represent the types in the terminology of [6]. The probability of the connections between any two vertices, say $u, v \in \mathbb{R}^d$, in these models is a function of the distance between them, e.g., it decays with the distance (e.g., [11], [8], or percolation models [14], [7]), or simply it can be a step function (e.g., [3]). These models are often designed to describe models of real world, as, e.g., various biological or social networks, where the frequency, or the probability of connections between two vertices often depend on the distance. Therefore, the analysis of such models often focuses on properties such as diameter ([3], [11]), cluster coefficient ([3]), propagation of a bootstrap percolation processes ([11], [8]).

First models of spread-out percolation (see [15] and [14], in particular) were studied even earlier than the theory in [6] was developed.

The relation between random graphs (on a finite set of vertices) and some spread-out percolation models (on a countable set of vertices) was exemplified in particular, by works [14] and [7]. In [7], the results of [14] on the critical parameters for the existence of an infinite cluster in a class of spread-out percolation where further generalized, and shown to be a direct consequence of the general theory of inhomogeneous random graphs [6].

Inspired by the models of inhomogeneous graphs [6] the authors of [9] generalized the model of [1] by introducing more inhomogeneity in the probabilities of connections (see details below).

Most of the investigations of the random distance graphs are being done without much application of the theory [6] (not counting the Example 4.6 in [6] itself, paper [7] is almost an exception). The reason, perhaps, is that although the random distance graphs are examples of a general model [6], they belong to the most complicated subclass of these models being outside of the so-called rank-1 case. Note that while the theory of [6] gives us the critical parameters for the emergence of the giant component and even describes the size of this component in the supercritical phase, there is no yet a general theory for the subcritical phase. The subcritical phase of the models which are not of rank 1 was studied only for some particular subclasses (see [18]), which do not include the present model. Furthermore, the critical phase was studied so far only for the rank-1 cases ([2], and later [4], [5], and [19]).

Here we consider a model recently introduced in [11]. As we demonstrate

below, it is “a simple homogeneous” case (as Example 4.6 in [6]). We generalize model of [11] in spirit of work [9]. This makes the model essentially inhomogeneous but still allows to get the exact asymptotics for the size of the largest connected component.

2 The Model

Let $N > 1$ and let V_N denote the set of vertices in the 2-dimensional discrete torus $\mathbb{T}_N^2 = (\mathbb{Z}/N\mathbb{Z})^2 = \{1, \dots, N\}^2$. Define the graph distance $d(u, v)$ between two vertices $u = (u_1, u_2)$ and $v = (v_1, v_2)$ in V_N as follows

$$d(u, v) = d_N(|u_1 - v_1|) + d_N(|u_2 - v_2|), \quad (2)$$

where

$$d_N(i) = \begin{cases} i, & 0 \leq i \leq N/2, \\ N - i, & N/2 < i < N, \end{cases}$$

for $i \in \{0, \dots, N - 1\}$.

Let W be a nonnegative random variable, and let W_v , $v \in V_N$, be *i.i.d.* copies of W . Given the values W_v , $v \in V_N$, assume that between any two vertices $u, v \in V_N$ an edge is present independently of others and with probability

$$p(u, v) = \min \left\{ c \frac{W_u W_v}{N d(u, v)}, 1 \right\},$$

where $c > 0$ is a parameter.

Denote $G_{N,W}$ the resulting graph on V_N .

In the case of the constant $W \equiv 1$ it is exactly the graph introduced and studied recently in [11]. This model is inspired by the study of neuronal networks (see references in [11]). It also has common features with other random graph models introduced previously in that area, see e.g., [17], where the same dynamics of activation was also approximated using the mean-field approach. On the other hand, this model is closely related to the models considered in [14], [7] and [9].

The work [11] focuses on a bootstrap percolation process, which in this case models the spread of the activation in a neuronal tissue. The authors of [11] also study the underlying graph and derive the order of its diameter, thus extending the results in [6] for the graph with unbounded number of types.

Here we describe the phase transition in the largest connected component. The relations between our random graph and the percolation model of [9] are exactly as shown in [7] for the homogeneous case of random distance graphs.

3 Results

It was disclosed already in [6] (Example 4.6) that in the supercritical case a homogeneous distance graph has the same asymptotics for the size of largest connected component as in the classical Erdős-Rényi model. We prove that this result holds for the subcritical case as well.

Theorem 1. *Let G_N denote a random graph on V_N with probability of connections*

$$p(u, v) = \min \left\{ \frac{c}{Nd(u, v)}, 1 \right\}, \quad u, v \in V_N,$$

and let C denote the size of the largest connected component in G_N . Set

$$\lambda = c 4 \log 2.$$

Then the following holds.

i) If $\lambda < 1$, we have that

$$\frac{C}{\log(N^2)} \xrightarrow{P} \frac{1}{\lambda - 1 - \log \lambda}, \quad \text{as } N \rightarrow \infty. \quad (3)$$

ii) If $\lambda > 1$ then

$$\frac{C}{N^2} \xrightarrow{P} \beta_\lambda, \quad \text{as } N \rightarrow \infty,$$

where $\beta = \beta(\lambda)$ is the positive solution of $\beta = 1 - e^{-\lambda\beta}$.

As we noted above, only the case *ii)* of this theorem follows by the results of [6].

Remark 1. *One may observe a certain redundancy here, as the statements *ii)* and *i)* of Theorem 1 are particular cases of the following below Theorems 2 and 3, respectively. However, stated separately Theorem 1 makes it clear that asymptotically the largest connected component in G_N behaves as the one in the Erdős-Rényi graph $G_{n,p}$ with $n = N^2$ and $p = \lambda/n$. Furthermore, it is plausible to conjecture (but we do not prove it here) that even in the critical case the model has the same asymptotics for the largest component as the one in $G_{n,p}$ with $p = 1/n$.*

The following theorem treats the general case (2).

Theorem 2. *Assume, that*

$$\mathbb{E}W^2 = \int_0^\infty x^2 \mu_W(dx) < \infty. \quad (4)$$

Let $C(G_{N,W})$ denote the size of its largest connected component in $G_{N,W}$, and denote again

$$\lambda = c 4 \log 2.$$

Then

$$\frac{C(G_{N,W})}{N^2} \xrightarrow{P} \int_0^\infty \beta(x) \mu_W(dx) =: \hat{\beta}, \quad (5)$$

where $\beta(x)$ is the maximal solution to

$$f(x) = 1 - e^{x\lambda \int_0^\infty y f(y) \mu_W(dy)}. \quad (6)$$

Furthermore, $\hat{\beta} > 0$ if and only if

$$\lambda \mathbb{E}W^2 > 1. \quad (7)$$

Note that the critical parameter $\lambda \mathbb{E}W^2$ in Theorem 2 is similar (in fact it has exactly the same meaning of a certain average of the degree of a vertex) to the lower bound derived in Theorem 4.1 in [9] to provide the necessary conditions for the percolation.

Theorem 2 follows essentially by the general theory of [6] as we explain below. It tells us that the asymptotics of the largest component in $G_{N,W}$ behaves as the one in the rank-1 random graph on V_N defined by the following probabilities of connections between any $u, v \in V_N$

$$p_1(u, v) = \min \left\{ \lambda \frac{W_u W_v}{N^2}, 1 \right\}. \quad (8)$$

In this case the largest connected component in the subcritical case is sensitive to the tail of the distribution of W . It is known that in the rank-1 case the size of the largest component varies between polynomial (as in [10]) and logarithmic (as in [16]) orders. This behaviour is projected to the model under consideration here. We shall state a particular case of logarithmic order to show the similarities with Theorem 1.

Theorem 3. *Assume that for some positive ε*

$$\mathbb{E}e^{\varepsilon W} < \infty. \quad (9)$$

If also

$$\lambda \mathbb{E}W^2 < 1, \quad (10)$$

there exists a unique $y > 1$ which satisfies

$$y = \frac{1}{\lambda M} \frac{\mathbb{E}(W e^{\lambda M(y-1)W})}{\mathbb{E}(W^2 e^{\lambda M(y-1)W})}, \quad (11)$$

where $M = \mathbb{E}W$. Let C be the size of the largest connected component of $G_{N,W}$. Then one has

$$\frac{C}{\log(N^2)} \xrightarrow{P} \frac{1}{\log \gamma}, \quad \text{as } N \rightarrow \infty,$$

where

$$\gamma := \frac{1}{\lambda \mathbb{E}(W^2 e^{\lambda M(y-1)W})} > 1.$$

4 Proofs

4.1 Random distance graph via inhomogeneous random graphs

Rescale the set V_N as follows

$$V_N \rightarrow \tilde{V}_N = \{(u_1/N, u_2/N) : (u_1, u_2) \in V_N\}.$$

Hence, \tilde{V}_N is a subset of N^2 vertices in a continuous torus $\mathbb{T}^2 := (\mathbb{R}/\mathbb{Z})^2$. Let $\mu_{\mathcal{L}}$ denote the Lebesgue measure on \mathbb{T}^2 , and let μ_W be the Borel measure on \mathbb{R}_+ induced by the random variable W . Denote $\mathcal{S} := \mathbb{T}^2 \times \mathbb{R}_+$, and define a product measure $\mu = \mu_{\mathcal{L}} \times \mu_W$ on this space. Then the triple $\mathcal{V} := (\mathcal{S}, \mu, \{(v, W_v) : v \in \tilde{V}_N\})$ satisfies the definition of a generalized vertex space of [6], i.e., for any Borel set $A \subseteq \mathcal{S}$ the following convergence holds:

$$\frac{|\{v : (v, W_v) \in A\}|}{N^2} \xrightarrow{P} \mu(A).$$

Define now for $u \neq v$, $u, v \in \mathbb{T}^2$ the kernel function

$$\kappa_1(u, v) := \frac{1}{\rho(u, v)}, \quad (12)$$

where for any $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{T}^2$

$$\rho(u, v) = \rho_1(|u_1 - v_1|) + \rho_1(|u_2 - v_2|)$$

with

$$\rho_1(a) = \begin{cases} a, & 0 \leq a \leq \frac{1}{2}, \\ 1 - a, & \frac{1}{2} < a \leq 1. \end{cases}$$

Denote also here $\kappa_2(x, y) := xy$ a product function on \mathbb{R}_+^2 .

Finally, using the kernel

$$\kappa((u, x), (v, y)) := \kappa_1(u, v)\kappa_2(x, y), \quad (13)$$

$(u, v), (x, y) \in \mathcal{S}$, we define a random graph $G^\mathcal{V}(N^2, \kappa)$ (we follow the notation of [6]) on the set of vertices \tilde{V}_N with given types

$$\tilde{\mathbf{V}}_N := \{(v, w_v), v \in \tilde{V}_N\} = \{(v_i, w_i)\}_{i=1}^{N^2} \subset \mathcal{S}$$

by setting an independent edge between any pair of vertices $\mathbf{x}_i, \mathbf{x}_j \in \tilde{\mathbf{V}}_N$ with probability (recall eq. (1))

$$\tilde{p}(\mathbf{x}_i, \mathbf{x}_j) := \min \left\{ c \frac{\kappa(\mathbf{x}_i, \mathbf{x}_j)}{N^2}, 1 \right\}. \quad (14)$$

Proposition 1. *The model G_N is equivalent to the inhomogeneous random graph model $G^\mathcal{V}(N^2, \kappa)$.*

Proof. Setting $\tilde{v} = v/N$ and $w_v = w_{\tilde{v}}$ for any $v \in V_N$, the probability of connection (2) is equivalent to

$$\begin{aligned} p(u, v) &= \min \left\{ c \frac{w_u w_v}{Nd(u, v)}, 1 \right\} = \min \left\{ c \frac{\kappa_1(\tilde{u}, \tilde{v}) w_{\tilde{u}} w_{\tilde{v}}}{N^2}, 1 \right\} \\ &= \tilde{p}((\tilde{v}, w_{\tilde{v}}), (\tilde{u}, w_{\tilde{u}})). \end{aligned} \quad (15)$$

Hence, there is a connection between any two vertices $u, v \in V_N$ of G_N , if and only if there is a connection between the corresponding vertices $\tilde{u} = u/N, \tilde{v} = v/N$ of $G^\mathcal{V}(N^2, \kappa)$. \square

It is straightforward to check that the kernel κ is *graphical*, since it is continuous, $\kappa \in L^1(\mathcal{S} \times \mathcal{S}, \mu \times \mu)$, and the number of edges in the graph $e(G^\mathcal{V}(N^2, \kappa))$ satisfies the following convergence

$$\frac{1}{N^2} \mathbb{E} e(G^\mathcal{V}(N^2, \kappa)) \rightarrow \frac{1}{2} \int_{\mathcal{S}} \int_{\mathcal{S}} \kappa(u, v) du dv. \quad (16)$$

4.2 Proof of Theorem 2

Proposition 1 together with (16) enables us to apply some of the results of [6] to our case. In particular, we can approximate the size of the connected component by the total progeny of a multi-type Galton-Watson branching process $\mathcal{B}(\mathbf{x})$, with type space \mathcal{S} , where the single ancestor has type \mathbf{x} , and

the number of offspring of type \mathbf{y} of each individual of type $\mathbf{x} \in \mathcal{S}$ has Poisson distribution with intensity $\kappa(\mathbf{x}, \mathbf{y})\mu(d\mathbf{y})$. Denote here $\beta_\kappa(\mathbf{x})$ and $\mathcal{X}(\mathbf{x})$, correspondingly, the survival probability and the size of the total progeny of this branching process with the ancestor of type \mathbf{x} .

Following [6], let us define the integral operator T_κ

$$(T_\kappa f)(\mathbf{x}) := \int_{\mathcal{S}} \kappa(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mu(\mathbf{y}),$$

for all measurable functions f (when the integral is defined) on \mathcal{S} , and define the norm of T_κ by

$$\|T_\kappa\| := \sup\{\|T_\kappa f\|_2 : f \geq 0, \|f\|_2 \leq 1\}. \quad (17)$$

Then, by Theorem 3.1 in [6] (whose applicability here is granted by Proposition 1) we immediately get

$$\frac{C(G_{N,W})}{N^2} \xrightarrow{P} \int_{\mathcal{S}} \beta_\kappa(\mathbf{x}) \mu(d\mathbf{x}). \quad (18)$$

Moreover, it is also proved in [6] that the survival probability β_κ is the maximal solution to

$$f(\mathbf{x}) = 1 - e^{-(T_\kappa f)(\mathbf{x})}, \quad \mathbf{x} \in \mathcal{S}, \quad (19)$$

and that $\hat{\beta}_\kappa > 0$ if and only if

$$\|T_\kappa\| > 1. \quad (20)$$

Observe that it follows directly from the symmetry of our model that the survival probability $\beta_\kappa(\mathbf{x})$ where $\mathbf{x} = (u, x) \in \mathbb{T}^2 \times \mathbb{R}_+$, does not depend on $u \in \mathbb{T}^2$, but it is only a function of $x \in \mathbb{R}_+$. Hence we shall simply write the survival probability as $\beta_\kappa(\mathbf{x}) = \beta_\kappa(x)$, which by (19) is the maximal solution to

$$f(x) = 1 - e^{-\lambda \int_0^\infty xy f(y) \mu_W(dy)}, \quad x \in \mathbb{R}_+, \quad (21)$$

which is exactly equation (6). This together with (18) yields (5).

We are left to prove (7). Firstly, one could use straightforwardly the definition (17) to derive

$$\|T_\kappa\| = \lambda \mathbb{E} W^2, \quad (22)$$

which together with (20) would give us (7). However, it is easier to derive (7) directly relating the defined above multi-type branching process to a certain homogeneous Galton-Watson process as it was explained, e.g., in [16].

Let us introduce yet another branching process $\mathcal{B}_1(x)$ with type space \mathbb{R}_+ , where the single ancestor has type x , and number of offspring of type $y \in \mathbb{R}_+$

of any individual of type $x \in \mathbb{R}_+$ has Poisson distribution with intensity $\lambda xy \mu_W(dy)$. By using the same analysis as for \mathcal{B} we can define the survival probability $\beta_\kappa^1(x)$ of $\mathcal{B}_1(x)$ as the maximum solution of (21). Therefore in notations of Theorem 2 it holds that $\beta(x) = \beta_\kappa^1(x)$, and

$$\beta_\kappa(\mathbf{x}) = \beta_\kappa^1(x) = \beta(x), \quad (23)$$

for any $\mathbf{x} = (u, x)$, and for all $u \in \mathbb{T}^2$.

Finally, we define one more auxiliary branching process \mathcal{B}_2 , this time a homogeneous Galton-Watson process. This process starts with one single ancestor, and its offspring distribution \tilde{Y} has a compound Poisson distribution

$$\text{Po}(\widetilde{W} \lambda \mathbb{E}(W)),$$

where random variable \widetilde{W} has the following so-called "size-biased", distribution:

$$\mu_{\widetilde{W}}(dy) := \frac{y \mu_W(dy)}{\mathbb{E} W}.$$

Let us denote $\mathcal{X}_1(x)$ and \mathcal{X}_2 the total progeny of $\mathcal{B}_1(x)$ and \mathcal{B}_2 , respectively. It is proved in Section 2.2 in [16], that $\mathcal{X}_1(\widetilde{W})$, and \mathcal{X}_2 are equal in distribution, i.e.,

$$\mathcal{X}_1(\widetilde{W}) \stackrel{d}{=} \mathcal{X}_2. \quad (24)$$

Hence, the critical parameters for the corresponding survival probabilities for these two branching processes $\mathcal{B}_1(\widetilde{W})$ and \mathcal{B}_2 are the same. In the case of a homogeneous process \mathcal{B}_2 the necessary and sufficient condition for a positive survival probability is simply $\mathbb{E}(\tilde{Y}) = \lambda \mathbb{E}(W^2) > 1$. Therefore (24) yields that $\mathbb{P}(\tilde{\mathcal{X}}_1(\widetilde{W}) = \infty) > 0$ if and only if $\lambda \mathbb{E}(W^2) > 1$.

It follows by the properties of Poisson distribution that the type of a randomly chosen offspring of the ancestor in the process $\mathcal{B}_1(x)$ has distribution \widetilde{W} . Hence, for any x the process $\mathcal{B}_1(x)$ survives with a positive probability, i.e., $\beta_\kappa^1(x) > 0$, if $\mathcal{B}_1(\widetilde{W})$ survives with a positive probability. Since β_κ^1 is the maximal solution to (21), i.e., to (6), it follows that $\hat{\beta} > 0$ (see (5)) in this case.

On the other hand, if $\mathcal{B}_1(\widetilde{W})$ survives with probability zero, then the equality

$$0 = \mathbb{P}(\tilde{\mathcal{X}}_1(\widetilde{W}) = \infty) = \int_{\mathbb{R}_+} \mathbb{P}(\tilde{\mathcal{X}}_1(x) = \infty) \mu_{\widetilde{W}}(dx) = \int_{\mathbb{R}_+} \beta_\kappa^1(x) \mu_{\widetilde{W}}(dx)$$

implies that $\beta_\kappa^1 = 0$ $\mu_{\widetilde{W}}$ -a.s., and hence, μ_W -a.s. Since β_κ^1 is the maximal solution to (21), i.e., to (6), it follows that in this case $\hat{\beta} = 0$.

Summarizing we get that $\hat{\beta} > 0$ if and only if $\lambda \mathbb{E}(W^2) > 1$. In turn this yields $\|T_\kappa\| = \lambda \mathbb{E}(W^2)$. This proves the theorem.

4.3 Proof of Theorem 1

4.3.1 Exploration algorithm.

Let v be an arbitrarily fixed vertex of a graph G , and let C_v denote the connected component containing v . We use a standard procedure to reveal C_v , called exploration algorithm. This is defined as follows.

In the course of exploration the vertices of G can be in one of the three states: *active*, *saturated* or *neutral*. At time $i = 0$ the vertex v is set to be active, while all the other vertices are set to be neutral. We denote by S_i the set of active vertices at time i . Hence, $|S_0| = 1$. The state of a vertex changes during the exploration of C_v as follows.

At each time step $i \geq 1$, we choose an active vertex in S_{i-1} uniformly at random and denote it by v_i . Then, consider each neutral vertex u : if u is connected to v_i then u becomes active, otherwise u stays neutral. After searching the entire set of neutral vertices the vertex v_i becomes saturated. This finishes the i -th step of the exploration algorithm.

The process stops when there are no more active vertices, i.e. at the first time i when $|S_i| = 0$, that is at time

$$T = \min\{i \geq 1 : |S_i| = 0\}. \quad (25)$$

At this time all considered vertices are saturated and they do not have any connection to the neutral vertices. Hence, C_v coincides with the set of saturated vertices, that is $|C_v| = T$.

Let X_i denote the number of vertices becoming active at the i -th step. Then the following recursion holds

$$\begin{aligned} |S_0| &= 1, \\ |S_i| &= |S_{i-1}| + X_i - 1 = X_1 + \cdots + X_i - (i - 1). \end{aligned} \quad (26)$$

Correspondingly, we can rewrite T , defined in eq. (25), as follows

$$T = \min\{i \geq 1 : X_1 + \cdots + X_i = i - 1\}. \quad (27)$$

4.3.2 Subcritical case

In this section we assume that $\lambda < 1$ and we prove part i) of Theorem 1.

Lower bound. We start by finding a lower bound on C , the size of the largest connected component. Namely, we prove that for any positive ε

$$\mathbb{P} \left\{ \frac{C}{\log N} \geq \frac{2}{1 - \lambda - \log \lambda} + \varepsilon \right\} \longrightarrow 0, \text{ as } N \longrightarrow \infty. \quad (28)$$

Proof. The proof is based on the exploration algorithm described above. We also use essentially the geometry of the discrete torus with the distance defined in eq. (2). Recall, in particular, that the number N_r of vertices at distance r from any given vertex, for N odd is given by

$$N_r = \begin{cases} 4r, & 1 \leq r \leq \lfloor N/2 \rfloor, \\ 4(N-r), & \lfloor N/2 \rfloor < r \leq N; \end{cases} \quad (29)$$

while for N even, it is given by

$$N_r = \begin{cases} 4r, & 1 \leq r < N/2, \\ 2(N-1), & r = \frac{N}{2}, \\ 4(N-r), & N/2 < r < N, \\ 1, & r = N. \end{cases} \quad (30)$$

Recall that the vertices becoming active at the i -th step are connected to the vertex v_i . Let $X_{i,r}$ denote the number of vertices at distance r from vertex v_i , which become active at the i -th step. Hence,

$$X_i = \sum_{r=1}^N X_{i,r}. \quad (31)$$

Let U_i denote the number of active and saturated vertices at time i (in other words, U_i is the number of vertices revealed by time i). In particular, by eq. (26) we have

$$U_i = |S_i| + i. \quad (32)$$

Correspondingly, for any vertex u let $U_{i,r}(u)$ be the number of active and saturated vertices at time i , which are at distance r from u . In particular, for any $i \geq 1$ and any vertex u it holds that

$$\sum_{r=1}^N U_{i,r}(u) = U_i.$$

The number $X_{i,r}$ depends on the number $U_{i-1,r}(v_i)$ of active and saturated vertices at time $i-1$ which are at distance r from v_i , in the following way

$$X_{i,r} | U_{i-1,r}(v_i) \in \text{Bin}(N_r - U_{i-1,r}(v_i), p_r), \quad (33)$$

where we use the notation

$$p_r = \frac{c}{N_r} = p(u, v), \quad \text{if } d(u, v) = r.$$

Remark 2. In eq. (33) and elsewhere we write a random parameter for a distribution with the usual meaning that the distribution is defined conditionally on a given value of the parameter.

Let us introduce the random variables

$$Z_{i,r} \in \text{Bin}(U_{i-1,r}(v_i), p_r),$$

and

$$X_{i,r}^+ = X_{i,r} + Z_{i,r} \in \text{Bin}(N_r, p_r).$$

Then, we define

$$X_i^+ := \sum_{r=1}^N X_{i,r}^+.$$

If a random variable ξ stochastically dominates η we denote this by $\eta \preceq \xi$.

It is clear from the above definition that $X_{i,r} \preceq X_{i,r}^+$, and, correspondingly, $X_i \preceq X_i^+$. Therefore,

$$|S_i| \preceq S_i^+ := X_1^+ + \cdots + X_i^+ - (i - 1).$$

Notice that the largest connected component has size larger than k if and only if there is a component of size at least k . Let $|C_v|$ denote the size of the component C_v . Then

$$\mathbb{P}\{C \geq k\} = \mathbb{P}\{\exists v : |C_v| \geq k\} = \mathbb{P}\left\{\bigcup_{v \in V} \{|C_v| \geq k\}\right\}.$$

It follows simply by the symmetry of the model that the random variables $|C_v|$ have equal distribution for all vertices v . This allows us to derive from the last formula the following bound

$$\mathbb{P}\{C \geq k\} \leq N^2 \mathbb{P}\{|C_v| \geq k\}, \quad (34)$$

for any arbitrarily fixed vertex v .

By the exploration algorithm we have that the probability for a component C_v to be larger or equal to k is equal to the probability of having active vertices in each of the $k - 1$ steps of the exploration, Hence,

$$\begin{aligned} \mathbb{P}\{|C_v| \geq k\} &= \mathbb{P}\{|S_t| > 0, \forall t \leq k - 1\} \\ &\leq \mathbb{P}\{S_t^+ > 0, \forall t \leq k - 1\} \\ &\leq \mathbb{P}\{S_{k-1}^+ > 0\} = \mathbb{P}\left\{\sum_{t=1}^{k-1} X_t^+ - (k - 2) > 0\right\}. \end{aligned} \quad (35)$$

We use the coupling method described in [12] for finding stochastic bounds on X_i^+ . We have that $X_{i,r}^+$ is stochastically bounded from above by a random variable $Y_{i,r} \stackrel{d}{=} \text{Po}(-N_r \log(1 - p_r))$, i.e. $X_{i,r}^+ \preceq Y_{i,r}$. Therefore, we can stochastically bound X_i^+ by a Poisson random variable, as follows

$$X_i^+ \preceq \sum_{r=1}^N Y_{i,r} \in \text{Po} \left(\sum_{r=1}^N -N_r \log(1 - p_r) \right) = \text{Po}(\lambda_N), \quad (36)$$

where

$$\begin{aligned} \lambda_N &= \sum_{r=1}^N -N_r \log(1 - p_r) = \sum_{r=1}^N N_r(p_r + o(p_r)) \\ &= \sum_{r=1}^{\lfloor N/2 \rfloor} 4r(p_r + o(p_r)) + \sum_{r=\lfloor N/2 \rfloor + 1}^N 4(N - r)(p_r + o(p_r)) \\ &= \lambda - \frac{2c}{N} + o\left(\frac{1}{N}\right). \end{aligned} \quad (37)$$

Let Y_i , $i \geq 1$, be the *i.i.d.* random variables with $\text{Po}(\lambda_N)$ distribution. Then we derive, using (35) and (36) with (37), the following upper bound for the probability in (34):

$$\mathbb{P}\{C \geq k\} \leq N^2 \mathbb{P} \left\{ \sum_{t=1}^{k-1} X_t^+ > k - 2 \right\} \leq N^2 \mathbb{P} \left\{ \sum_{t=1}^{k-1} Y_t > k - 2 \right\}. \quad (38)$$

Using Chebyshev's inequality in (38), for any $h > 0$, we have that

$$\begin{aligned} \mathbb{P}\{C \geq k\} &\leq N^2 \mathbb{P} \left\{ \sum_{t=1}^{k-1} Y_t > k - 2 \right\} \\ &\leq \frac{N^2 \prod_{t=1}^{k-1} \mathbb{E} e^{hY_t}}{e^{h(k-2)}} \\ &= N^2 e^{-h(k-2)} \prod_{t=1}^{k-1} e^{\lambda_N(e^h - 1)} \\ &= N^2 e^{-h(k-2)} e^{(k-1)\lambda_N(e^h - 1)}. \end{aligned} \quad (39)$$

The last formula attains its minimum at $h = \log\left(\frac{k-1}{k\lambda}\right)$, where it equals to

$$N^2 e^{k(1 - \lambda + \log \lambda) + k o(1)}.$$

Therefore setting $k = \left(\frac{2}{\lambda - 1 - \log \lambda} + \varepsilon\right) \log N$ in (39), we find that (28) holds. \square

The upper bound. To complete the proof of (3) we will prove that for any $\varepsilon > 0$

$$\mathbb{P} \left\{ \frac{C}{\log N} \leq \frac{2}{1 - \lambda - \log \lambda} - \varepsilon \right\} \longrightarrow 0 \quad \text{as } N \longrightarrow \infty. \quad (40)$$

Proof. Before proceeding to the proof of (40) we derive a useful result, which roughly speaking tells us that removing an arbitrary set of $o(N^2)$ vertices from V_N does not change some, asymptotically as $N \rightarrow \infty$, the expected degree of a vertex.

Lemma 1. *Let n_r , $r = 1, \dots, N$, with $0 \leq n_r \leq N_r$, be an arbitrary sequence such that*

$$\sum_{r=1}^N n_r = o(N^2).$$

Then

$$\frac{1}{N} \sum_{r=1}^N \frac{n_r}{r} \rightarrow 0 \quad \text{as } N \rightarrow \infty. \quad (41)$$

Proof. We prove the lemma by contradiction. Assume there exists a constant $c > 0$ such that for any $M \in \mathbb{N}$ there exists $N \geq M$ such that

$$\frac{1}{N} \sum_{r=1}^N \frac{n_r}{r} \geq c. \quad (42)$$

Let $0 < \delta < \min\{4, c\}$ and define the sets \mathcal{N}_δ and its complementary $\overline{\mathcal{N}}_\delta$ as follows

$$\begin{aligned} \mathcal{N}_\delta &= \{r \in \{1, \dots, N\} : n_r \geq \delta r\} \\ \overline{\mathcal{N}}_\delta &= \{r \in \{1, \dots, N\} : n_r < \delta r\}. \end{aligned}$$

Noting that from (29)-(30) we have $n_r \leq N_r \leq 4r$, for any $0 \leq r \leq N$, from (42) it follows that

$$\begin{aligned} c &\leq \frac{1}{N} \sum_{r=1}^N \frac{n_r}{r} = \frac{1}{N} \left(\sum_{r \in \mathcal{N}_\delta} \frac{n_r}{r} + \sum_{r \in \overline{\mathcal{N}}_\delta} \frac{n_r}{r} \right) \\ &\leq \frac{1}{N} \left(\sum_{r \in \mathcal{N}_\delta} 4 + \sum_{r \in \overline{\mathcal{N}}_\delta} \delta \right) \\ &= \frac{1}{N} (4|\mathcal{N}_\delta| + \delta \overline{\mathcal{N}}_\delta) \\ &= \delta + \frac{4 - \delta}{N} |\mathcal{N}_\delta|. \end{aligned}$$

In particular, we have

$$|N_\delta| \geq \frac{c-\delta}{4-\delta}N,$$

and therefore

$$\sum_{r=1}^N n_r \geq \sum_{r \in \mathcal{N}_\delta} n_r \geq \sum_{r \in \mathcal{N}_\delta} \delta r \geq \delta \sum_{r=1}^{|\mathcal{N}_\delta|} r \geq \frac{\delta}{2} |\mathcal{N}_\delta|^2 \geq \frac{\delta}{2} \left(\frac{c-\delta}{4-\delta} \right)^2 N^2,$$

which contradicts the assumptions. \square

Now we can prove (40). We shall follow the construction used already in [16]. For any vertex v let $V(C_v)$ denote here the set of vertices of the component C_v . Observe that G_N can be decomposed into pairwise disjoint connected components as follows. Set $\tilde{v}_1 = v$. Then, given $C_{\tilde{v}_1}, \dots, C_{\tilde{v}_k}$, for $k \geq 1$ choose a vertex \tilde{v}_{k+1} uniformly in $V_N \setminus \cup_{i=1}^k V(C_{\tilde{v}_i})$, unless the last set is empty, in which case we stop the algorithm. The graph G_N is thus decomposed into pairwise disjoint components $C_{\tilde{v}_1}, \dots, C_{\tilde{v}_M}$, with $M = M(N)$. Note that here M is a bounded random variable, $1 \leq M \leq N^2$, denoting the number of disjoint connect components in G_N .

Fix $\varepsilon > 0$ arbitrarily and denote $K = \left(\frac{2}{\lambda-1-\log \lambda} + \varepsilon \right) \log N$. Then we define the event

$$E_N = \{C \leq K\}.$$

Recall that it follows from (28) that

$$\mathbb{P}\{\bar{E}_N\} \rightarrow 0$$

as $N \rightarrow \infty$. This yields

$$\begin{aligned} \mathbb{P}\{C \leq k\} &= \mathbb{P}\{|C_{\tilde{v}_1}| \leq k, \dots, |C_{\tilde{v}_M}| \leq k\} \\ &\leq \mathbb{P}\{|C_{\tilde{v}_1}| \leq k, \dots, |C_{\tilde{v}_M}| \leq k | E_N\} + o(1). \end{aligned} \tag{43}$$

Note that since conditionally on E_N the largest connected component is smaller than K , it follows that $MK \geq N^2$. Hence, for any

$$M_1 \leq \frac{N^2}{K} \leq M$$

the following bound holds from the probability in (43):

$$\begin{aligned} &\mathbb{P}\{|C_{\tilde{v}_1}| \leq k, \dots, |C_{\tilde{v}_M}| \leq k | E_N\} \\ &\leq \prod_{i=1}^{M_1} \mathbb{P}\{|C_{\tilde{v}_i}| \leq k | |C_{\tilde{v}_1}| \leq k, \dots, |C_{\tilde{v}_{i-1}}| \leq k, E_N\}. \end{aligned} \tag{44}$$

Let \mathcal{V}_0 be an arbitrary set of $M_1 K$ nodes, u be a vertex in $V_N \setminus \mathcal{V}_0$, and let $\tilde{C}_u = \tilde{C}_u(\mathcal{V}_0)$ denote the connected component containing u constructed precisely as the original C_v but on the vertex set $V_N \setminus \mathcal{V}_0$.

Then, each factor in eq. (44) can be uniformly bounded as follows

$$\mathbb{P}\{|C_{\tilde{v}_i}| \leq k | |C_{\tilde{v}_1}| \leq k, \dots, |C_{\tilde{v}_{i-1}}| \leq k, E_N\} \leq \max_{\substack{\mathcal{V}_0 \subseteq V_N: |\mathcal{V}_0| = M_1 K \\ u \in V_N \setminus \mathcal{V}_0}} \mathbb{P}\{|\tilde{C}_u| \leq k\},$$

where we simply used the fact that on a smaller set of vertices the components are smaller as well. Therefore, from (44) it follows that

$$\mathbb{P}\{C \leq k\} \leq \left(\max_{\substack{\mathcal{V}_0: |\mathcal{V}_0| = M_1 K \\ u \in V_N \setminus \mathcal{V}_0}} \mathbb{P}\{|\tilde{C}_u| \leq k\} \right)^{M_1}. \quad (45)$$

In the following, fix set \mathcal{V}_0 arbitrarily but so that $|\mathcal{V}_0| = M_1 K = o(N^2)$. Fix a vertex $u \notin \mathcal{V}_0$ arbitrarily, and construct the component \tilde{C}_u on the vertex set $V_N \setminus \mathcal{V}_0$ as it is described in the exploration algorithm. Let us denote here u_1, u_2, \dots , the sequence of saturated vertices.

Define $n_r^0(u)$ to be the number of nodes in \mathcal{V}_0 which are at distance r from u , so that $0 \leq n_r^0(u) \leq N_r$ and $\sum_{r=1}^N n_r^0(u) = |\mathcal{V}_0|$ for any u .

In analogy with the notion used previously, let here \tilde{U}_i denote the number of active and saturated vertices at step i in the exploration process (cfr. eq. (32)), and $\tilde{U}_{i,r}(w)$ be the number of those vertices at distance r from the vertex w . Let then $n_{i,r}^0 = n_r^0(u_i)$ denote the number of those vertices in \mathcal{V}_0 which are at distance r from the i -th saturated vertex u_i . By this definition, and our assumption on $|\mathcal{V}_0| = o(N^2)$ we have

$$\sum_{r=1}^N n_{i,r}^0 = |\mathcal{V}_0| = M_1 K = o(N^2) \quad (46)$$

for any i . Hence, the number of vertices activated at step i at distance r from the i -th explored vertex, which we denote $\tilde{X}_{i,r}$, has the following distribution

$$\tilde{X}_{i,r} \in \text{Bin}(N_r - n_{i,r}^0 - \tilde{U}_{i-1,r}(u_i), p_r),$$

and the total number of vertices activated at the i -th step is given by

$$\tilde{X}_i = \sum_{r=1}^N \tilde{X}_{i,r}.$$

Using these definitions we derive

$$\mathbb{P}\{|\tilde{C}_u| > k\} \geq \mathbb{P}\{\tilde{X}_1 + \tilde{X}_2 + \dots + \tilde{X}_t > t - 1, \forall t \leq k - 1\}. \quad (47)$$

To approximate the distribution of \tilde{X}_i is the last formula, let us define, given numbers $0 \leq k_{i,r} \leq N_r - n_{i,r}^0$, $r = 1, \dots, N$, $i = 1, \dots, k-1$, such that

$$\sum_{r=1}^N k_{i,r} \leq k, \quad (48)$$

the following Poisson random variables

$$\tilde{Z}_{i,r} \in \text{Po}((N_r - n_{i,r}^0 - k_{i,r})p_r).$$

Then

$$\tilde{Z}_i = \sum_{r=1}^N \tilde{Z}_{i,r} \in \text{Po}(\lambda_{i,N}), \quad (49)$$

where

$$\lambda_{i,N} = \sum_{r=1}^N (N_r - n_{i,r}^0 - k_{i,r})p_r. \quad (50)$$

To simplify the notations define the event

$$F_N = \{\tilde{U}_{i,r} = k_{i,r}, \quad \sum_{r=1}^N k_{i,r} \leq k, \quad \forall i \leq k-1, \forall r = 1, \dots, N\} \quad (51)$$

and consider

$$\begin{aligned} \mathbb{P}(\tilde{Z}_i > k | F_N) &= \mathbb{P}(\tilde{Z}_i > k, \tilde{Z}_i = \tilde{X}_i | F_N) \\ &\quad + \mathbb{P}(\tilde{Z}_i > k, \tilde{Z}_i \neq \tilde{X}_i | F_N) \\ &\leq \mathbb{P}(\tilde{X}_i > k | F_N) + \mathbb{P}(\tilde{Z}_i \neq \tilde{X}_i | F_N). \end{aligned}$$

Note that

$$\begin{aligned} \mathbb{P}\{\tilde{X}_i \neq \tilde{Z}_i | F_N\} &= \mathbb{P}\left\{\sum_{r=1}^N \tilde{X}_{i,r} \neq \sum_{r=1}^N \tilde{Z}_{i,r} | F_N\right\} \\ &\leq \mathbb{P}\left\{\bigcup_{r=1}^N \{\tilde{X}_{i,r} \neq \tilde{Z}_{i,r}\} | F_N\right\} \\ &\leq \sum_{r=1}^N \mathbb{P}\{\tilde{X}_{i,r} \neq \tilde{Z}_{i,r} | F_N\}. \end{aligned}$$

We shall use the following result from [20].

Lemma 2. [20] (Poisson limit for binomial random variable) $X \stackrel{d}{=} \text{Bin}(n, \lambda/n)$, $Y \stackrel{d}{=} \text{Po}(\lambda)$ then the following holds

$$\mathbb{P}(X \neq Y) \leq \frac{\lambda^2}{n}.$$

By Lemma 2 we have

$$\mathbb{P}\{\tilde{X}_{i,r} \neq \tilde{Z}_{i,r} | F_N\} \leq p_r^2(N_r - n_{i,r}^0 - k_{i,r}),$$

which gives us

$$\begin{aligned} \mathbb{P}\{\tilde{X}_i \neq \tilde{Z}_i | F_N\} &\leq \sum_{r=1}^N p_r^2(N_r - n_{i,r}^0 - k_{i,r}) \\ &= \frac{c^2}{N^2} \sum_{r=1}^N \frac{1}{r^2}(N_r - n_{i,r}^0 - k_{i,r}) = O\left(\frac{\log N}{N^2}\right), \end{aligned}$$

uniformly in i .

Next we consider

$$\begin{aligned} \mathbb{P}\{\tilde{Z}_1 + \dots + \tilde{Z}_t > t - 1, \forall t \leq k - 1\} \\ \leq \mathbb{P}\{\tilde{X}_1 + \dots + \tilde{X}_t > t - 1, \forall t \leq k - 1 | F_N\} \\ + \sum_{s=1}^{k-1} \mathbb{P}\{\tilde{X}_s \neq \tilde{Z}_s | F_N\}. \end{aligned}$$

Note that by (4.3.2)

$$\varepsilon_k(N) := \sum_{s=1}^{k-1} \mathbb{P}\{\tilde{X}_s \neq \tilde{Z}_s | F_N\} = O\left(\left(\frac{\log N}{N}\right)^2\right), \quad (52)$$

therefore (4.3.2) yields

$$\begin{aligned} \mathbb{P}\{\tilde{X}_1 + \dots + \tilde{X}_t > t - 1, \forall t \leq k - 1 | F_N\} \\ \geq \mathbb{P}\{\tilde{Z}_1 + \dots + \tilde{Z}_t > t - 1, \forall t \leq k - 1\} - \varepsilon_k(N). \end{aligned} \quad (53)$$

In the following we construct a random variable \tilde{Z}^- is stochastically smaller than any of the random variables \tilde{Z}_i , $k \leq i \leq k$.

First, using the assumption (46) and Lemma 1 we derive

$$\sum_r n_{i,r}^0 p_r = \frac{c}{N} \sum_r \frac{n_{i,r}^0}{r} = o(1). \quad (54)$$

From now on we shall assume that

$$k = a \log N \quad (55)$$

for some positive a . Under this assumption we have

$$\sum_r k_{i,r} p_r = \frac{c}{N} \sum_{r=1}^N \frac{k_{i,r}}{r} \leq \frac{c}{N} \sum_r k_{i,r} = \frac{kc}{N} = \frac{ac \log N}{N} = o(1). \quad (56)$$

Hence, from (54) and (56) we obtain the following bound for $\lambda_{i,N}$ defined in (50)

$$\lambda_{i,N} = \sum_{r=1}^N (N_r - n_{i,r}^0 - k_{i,r}) p_r \geq \sum_{r=1}^N N_r p_r + o_i(1), \quad (57)$$

where $o_i(1)$ might depend on i . Note that by (37)

$$\sum_{r=1}^N N_r p_r = \lambda + o(1). \quad (58)$$

Hence, for any (constant)

$$\lambda' < \lambda \quad (59)$$

(57) together with (58) yields the following uniform in $i \leq k$ bound

$$\lambda_{i,N} > \lambda'. \quad (60)$$

Recall that $\tilde{Z}_i \in \text{Po}(\lambda_{i,N})$ by (49). Therefore (60) allows us to construct independent $\tilde{Z}_i^- \in \text{Po}(\lambda')$, $1 \leq i \leq k$, such that

$$\tilde{Z}_i^- \leq \tilde{Z}_i \quad (61)$$

a.s. for each i . Now we can derive the following bound

$$\begin{aligned} \mathbb{P}(\tilde{Z}_1 + \dots, \tilde{Z}_t > t-1, t=1, \dots, k-1) \\ \geq \mathbb{P}(\tilde{Z}_1^- + \dots + \tilde{Z}_t^- > t-1, t=1, \dots, k-1) \\ = \mathbb{P}\{\mathcal{T} \geq k\}, \end{aligned} \quad (62)$$

where \mathcal{T} denotes the total progeny of a branching process with offspring distribution $\text{Po}(\lambda')$. Substituting this bound into (53) we obtain

$$\mathbb{P}\{\tilde{X}_1 + \dots + \tilde{X}_t > t-1, \forall t \leq k-1 | F_N\} \geq \mathbb{P}\{\mathcal{T} \geq k\} - \varepsilon_k(N), \quad (63)$$

where the right-hand side is uniform in F_N (where we assume only conditions in eqs. (48) and (55)). This yields

$$\mathbb{P}\{\tilde{X}_1 + \dots + \tilde{X}_t > t - 1, \forall t \leq k - 1\} \geq \mathbb{P}\{\mathcal{T} \geq k\} - \varepsilon_k(N),$$

and therefore, by eq. (47),

$$\mathbb{P}\{|\tilde{C}_u| \leq k\} \leq 1 - \mathbb{P}\{\mathcal{T} \geq k\} + \varepsilon_k(N). \quad (64)$$

Using a well known formula for the distribution of the progeny of a branching process (see e.g. [13]), we compute

$$\mathbb{P}\{\mathcal{T} \geq k\} = \sum_{j=k}^{\infty} \frac{(\lambda' j)^{j-1}}{j!} e^{-\lambda' j} \geq \frac{(\lambda' k)^{k-1}}{k!} e^{-\lambda' k},$$

which, together with the Stirling formula, gives us

$$\mathbb{P}\{\mathcal{T} \geq k\} \geq \frac{1}{\sqrt{2\pi\lambda'}} \frac{1}{k^{3/2}} e^{-\alpha k} \left(1 + O\left(\frac{1}{k}\right)\right), \quad (65)$$

where

$$\alpha = \lambda' - 1 - \log \lambda'. \quad (66)$$

Substituting eq. (65) into eq. (64), we obtain for $k = a \log N$ that

$$\mathbb{P}\{|\tilde{C}_u| \leq a \log N\} \leq 1 - \frac{1}{A_N} (1 + o(1)) + O\left(\left(\frac{\log N}{N}\right)^2\right), \quad (67)$$

where

$$A_N = \sqrt{2\pi\lambda'} (a \log N)^{3/2} N^{a\alpha}. \quad (68)$$

Choose now arbitrarily a constant

$$a < \frac{2}{\alpha}. \quad (69)$$

Then, eq. (67) yields

$$\mathbb{P}\{|\tilde{C}_u| \leq a \log N\} \leq 1 - \frac{1}{A_N} (1 + o(1)). \quad (70)$$

Observe that the value on the right in the last formula is uniform in the choice of the set \mathcal{V}_0 and vertex u . Therefore we can use the bound in eq. (70) to obtain, from eq. (45),

$$\mathbb{P}\{C \leq a \log N\} \leq \left(1 - \frac{1}{A_N} (1 + o(1))\right)^{M_1}. \quad (71)$$

Finally, we choose

$$M_1 = A_N \log N \gg A_N,$$

which by (69) satisfies as well the condition (46), i.e.,

$$M_1 K = o(N^2),$$

where

$$K = \left(\frac{2}{\lambda - 1 - \log \lambda} + \varepsilon \right) \log N.$$

With this choice of M_1 bound (71) implies

$$\mathbb{P}\{C \leq a \log N\} = o(1) \tag{72}$$

for any fixed constant (see eqs(69) and (66))

$$a < \frac{2}{\alpha} = \frac{2}{\lambda' - 1 - \log \lambda'}.$$

Since by (59) here we can choose any $\lambda' < \lambda$, it follows that (72) holds for any

$$a < \frac{2}{\lambda - 1 - \log \lambda}.$$

This proves (40), and therefore part *i*) of Theorem 1 is proved. \square

This completes the proof of Theorem 1, since part *ii*) follows by Theorem 2.

4.4 The idea of proof of Theorem 3

A proof of Theorem 3 can run precisely the lines of the proof of part *i*) of Theorem 1 in a combination with the proof of a corresponding result for the rank-1 model (8) given in [16]. Therefore, we can omit the details here.

References

- [1] M. Aizenman and C. M. Newman. Discontinuity of the percolation density in one dimensional $1/|x - y|^2$ percolation models. *Comm. Math. Phys.*, 107:611–647, 1986.
- [2] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25: 812–854, 1997.

- [3] C. Avin. Distance graphs: from random geometric graphs to Bernoulli graphs and between. In *DIALM-POMC*, 2008.
- [4] S. Bhamidi, R. van der Hofstad, and J. van Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electron. J. Probab.*, 15: 1682–1702, 2010.
- [5] S. Bhamidi, R. van der Hofstad and J. van Leeuwaarden. Novel scaling limits for critical inhomogeneous random graphs. *Ann. Probab.*, 40: 2299–2361, 2012.
- [6] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Algor.*, 31:3–122, 2007.
- [7] B. Bollobás, S. Janson, and O. Riordan. Spread-out percolation in \mathbb{R}^d . *Random Struct. Algor.*, 31:239–246, 2007.
- [8] M. Bradonjić and I. Saniee. Bootstrap percolation on random geometric graphs. *Probab. Eng. Inform. Sc.*, 28:169–181, 2014.
- [9] M. Deijfen, R. van der Hofstad, and G. Hooghiemstra. Scale-free percolation. *Ann. I. H. Poincaré-PR*, 49:817–838, 2013.
- [10] S. Janson. The largest component in a subcritical random graph with a power law degree distribution. *Ann. Appl. Probab.*, 18:1651–1668, 2008.
- [11] S. Janson, R. Kozma, M. Ruzinkó, and Y. Sokolov. Bootstrap percolation on a random graph coupled with a lattice. arXiv:1507.07997v2, 2015.
- [12] T. Lindvall. *Lectures on the coupling method*. Wiley-Interscience Publication, 1992.
- [13] R. Otter. The multiplicative process. *Ann. Math. Statist.*, 20:206–224, 1949.
- [14] M. Penrose. On the spread-out limit for bond and continuum percolation. *Ann. Appl. Probab.*, 3:253–276, 1993.
- [15] M. Penrose. *Random geometric graphs*. Oxford Studies in Probability, 5. Oxford University Press, 2003.
- [16] T. S. Turova. The size of the largest component below phase transition in inhomogeneous random graphs. *Comb. Probab. Comp.*, 20:131–154, 2010.

- [17] T. S. Turova. The emergence of connectivity in neuronal networks: from bootstrap percolation to auto-associative memory. *Brain Res.*, 1434:277–284, 2012.
- [18] T. S. Turova. Asymptotics for the size of the largest component scaled to “log n” in inhomogeneous random graphs. *Ark. Mat.*, 51:371–403, 2013.
- [19] T. S. Turova. Diffusion approximation for the components in critical inhomogeneous random graphs of rank 1. *Random Struct. Alg.*, 43:486 – 539, 2013.
- [20] R. van der Hofstad. Random graphs and complex networks, Vol. I. <http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf>, 2016.